

Enquête mensuelle sur les services de restauration et débits de boissons (EMSRDB) : répercussions de la restructuration

Le présent rapport décrit la restructuration de la base de sondage de l'EMSRDB et la préparation d'un échantillon à jour. Il explique comment les anciennes séries mensuelles de l'EMSRDB (avant la restructuration) et les nouvelles séries (après la restructuration) ont été reliées.

La restructuration a pour but d'améliorer l'échantillon en prenant en considération les changements qui se produisent dans une population au fil du temps. Le plan d'échantillonnage initial de l'EMSRDB était fondé sur la liste des entreprises et des établissements inscrits au Registre des entreprises (RE) au mois de mars 2007. L'échantillonnage restructuré est fondé sur la liste du RE de décembre 2008.

Les deux plans ont été utilisés parallèlement pendant six mois aux fins de surveillance de la qualité et, plus important encore, afin de trouver un moyen d'établir des liens entre les nouvelles et les anciennes séries. Cela a permis d'élaborer une série révisée couvrant la période de septembre 2007 à mars 2009. La nouvelle série restructurée est en production de façon indépendante depuis avril 2009.

La section 1 du présent document traite des événements qui ont mené à la nécessité de procéder à une restructuration. La section 2 compare les estimations des anciennes et des nouvelles séries. La section 3 décrit la méthodologie utilisée pour établir des liens entre les anciennes et les nouvelles séries.

L'annexe 1 contient un bref glossaire des termes et acronymes. L'annexe 2 donne un aperçu de l'historique et de la méthodologie de l'EMSRDB. Il sera utile de se reporter aux annexes pendant les sections ci-dessous.

Section 1 : La nécessité d'une restructuration

L'échantillon initial, c'est-à-dire avant la restructuration, était basé sur la liste des établissements du Registre des entreprises ayant le code du Système de classification des industries de l'Amérique du Nord (SCIAN) 722 au mois de mars 2007 (la population initiale de l'EMSRDB). La liste des établissements, inscrits dans le RE, change tous les mois. Il y a des naissances et des fermetures. Des établissements deviennent plus petits ou plus grands et changent parfois de province ou d'industrie. Par conséquent, plus le temps passe, moins les échantillons sont optimaux.

La stratégie du plan d'échantillonnage de l'EMSRDB comporte des processus visant à garantir l'exactitude et la grande qualité des estimations malgré les populations qui changent constamment. Cependant, il est inévitable que l'erreur type des estimations de l'EMSRDB augmente. Les processus sont les suivants:

- Quand des fermetures sont détectées dans l'échantillon de l'EMSRDB, les entreprises ne sont pas enlevées de l'échantillon. Ces fermetures ne contribuent

plus aux estimations des ventes. Elles représentent toutefois les fermetures non détectées dans l'ensemble de la population de l'ESMRDB.

- Quand des entreprises naissent (tous les mois dans le RE), on les ajoute à la population admissible de l'ESMRDB et on sélectionne un échantillon.
- Lorsqu'une unité échantillonnée reste dans la population admissible, mais change de province ou de numéro SCIAN admissible initial, elle continue de contribuer. Toutefois, sa contribution va à sa nouvelle province ou à son nouveau SCIAN.

Malgré ces ajustements, au fil des mois, l'échantillon initial devient périmé et n'est plus optimal. Un échantillon non optimal aboutit à une erreur type plus grande (ou c.v.) que l'échantillon initial et fournit des estimations moins fiables. En outre, il est possible qu'un biais inconnu se glisse dans les estimations si la structure de la population change beaucoup depuis l'échantillon initial.

Pour ces raisons, la base de sondage de l'ESMRDB est restratifiée et un nouvel échantillon est sélectionné. La stratégie de l'ESMRDB consiste à procéder à une restratification et à sélectionner un nouvel échantillon tous les deux ans pour le mois de référence de décembre. Nous nous fonderons sur les résultats de la restratification terminée récemment pour évaluer les avantages et le moment des restratifications futures de l'ESMRDB.

En comparant les unités d'échantillonnage de la base de sondage initiale et de la nouvelle base de sondage de l'ESMRDB : 1,2 % des unités avaient une province ou un code SCIAN différents; 23 % étaient des nouvelles unités (naissances) et 24 % des fermetures. De plus, 17 % étaient passés dans une strate différente du plan d'échantillonnage. Globalement, cela représente un nombre important de changements dans la population de l'ESMRDB et le plan d'échantillonnage.

Durant les 21 mois qui se sont écoulés entre le nouveau et l'ancien plan d'échantillonnage, deux changements importants se sont produits :

- Un nombre considérable de nouveaux contrats de franchise ont été conclus et intégrés à la collecte des données. Un contrat de franchise est une entente conclue avec un franchiseur pour qu'il fournisse des données regroupées par province et par code SCIAN pour lui-même et ses franchisés.

Dans l'ancien échantillon, il y avait 25 contrats de franchise. Ce nombre a augmenté à 47 pour le nouvel échantillon.

Les contrats de franchise aident à améliorer l'efficacité de la collecte des données, à réduire le fardeau de réponse pour les franchisés et à diminuer la variabilité de la taille de l'échantillon, étant donné que davantage d'unités de population sont couvertes sans augmenter la taille de l'échantillon. Ils contribuent aussi à garantir

une couverture de la population plus complète et à jour, car les fermetures, les naissances et les unités manquantes de la population sont saisies immédiatement.

- On a terminé le projet de refonte du RE à temps pour le mois de référence de janvier 2008. Un aspect important de la refonte du RE était une mesure améliorée de la taille des entreprises. La mesure de la taille est une variable essentielle utilisée dans le plan d'échantillonnage, ce qui suppose un plan d'échantillonnage plus efficace pour la restratification.

Le diagramme 5 (se reporter à l'annexe 2) illustre le plan d'échantillonnage. Cela s'applique à la fois au plan initial et au plan restratifié. Le principal objectif d'une restratification est de calculer les limites révisées et optimales du plan d'échantillonnage entre les strates à tirage complet, à tirage partiel et à tirage nul (se reporter à l'annexe 1). Les strates mises à jour reflètent mieux la population réelle de l'ESMRDB.

Dans la restratification de l'ESMRDB, nous avons choisi l'échantillon indépendamment de l'échantillon précédent. Cependant, il y a un chevauchement important entre l'ancien et le nouvel échantillon étant donné que les unités d'échantillonnage les plus importantes, attribuées aux strates à tirage complet initiales, resteront probablement importantes et seront à nouveau attribuées aux strates à tirage complet dans le nouvel échantillon.

Section 2 : Comparaison des anciennes et des nouvelles estimations de l'EMSRDB

L'échantillon initial (appelé C1) a été utilisé parallèlement au nouvel échantillon de restratification (appelé C2). Il y a deux ensembles d'estimations pour la période allant de décembre 2008 à mai 2009.

Les estimations de l'échantillon C2 devraient être plus exactes parce qu'elles sont fondées sur les renseignements les plus récents disponibles dans le RE. En outre, des biais inconnus ou ne pouvant être mesurés peuvent s'être glissés dans les estimations de l'échantillon C1, car la structure de la population de l'EMSRDB a changé depuis la création de la base de sondage initiale.

i) Variations mensuelles

Les variations mensuelles des estimations des échantillons C1 et C2 sont très similaires. Ceci est plus évident lorsque les petits domaines sont fusionnés pour obtenir des domaines plus importants. Pour cette raison, en établissant des liens entre les séries initiales C1 et les nouvelles séries C2, nous pourrions choisir un échantillon ou l'autre pour fournir les mouvements mensuels.

ii) Écarts de niveau

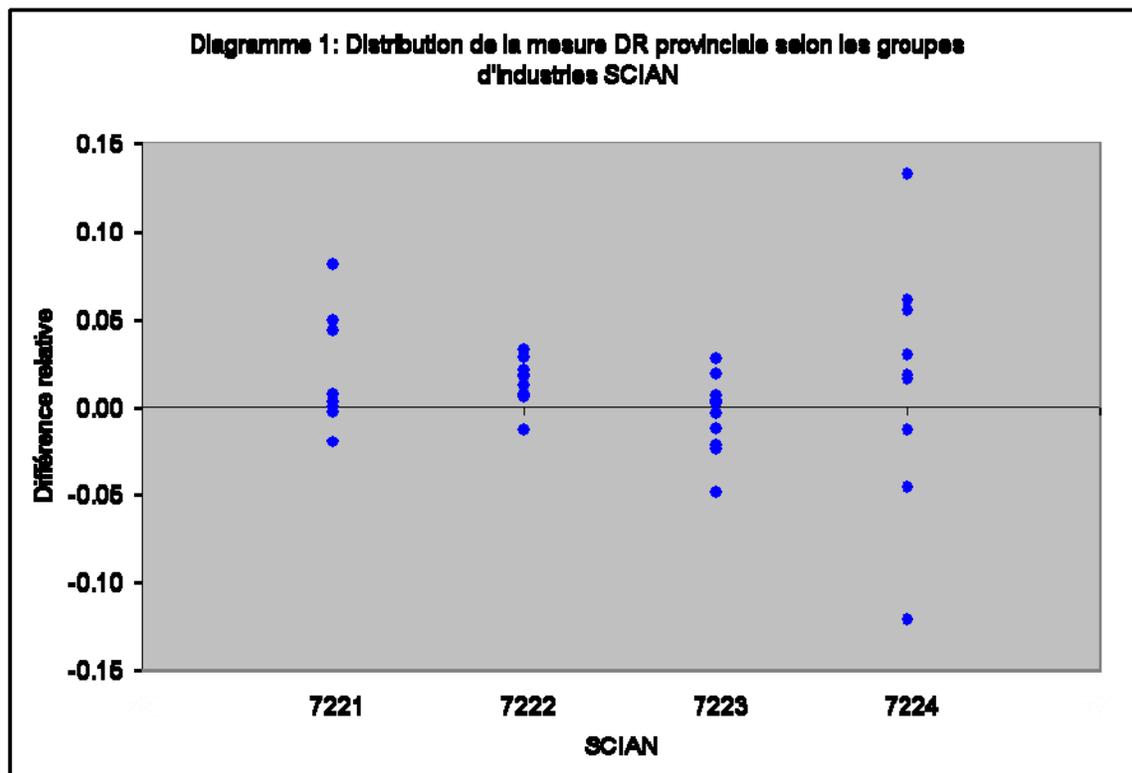
Le diagramme 1 représente quatre graphiques à points comparant les estimations C1 et C2 pour le mois de mars 2008 à l'aide de la différence relative : $DR = (C2 - C1)/(C2 + C1)$. La mesure DR est symétrique par rapport à 0 et se situe entre -1 et +1. Chaque

point représente la DR entre C1 et C2 pour une province particulière pour les groupes du SCIAN qui y correspondent.

Pour les groupes du SCIAN 7221 et 7223, la majorité des points forme une grappe autour de 0, avec une répartition relativement petite. Cela aboutit à la conclusion, qu'en moyenne, l'écart entre C1 et C2 pour ces deux classes du SCIAN est très petit.

Pour le groupe du SCIAN 7222, tous les points sauf un, sont au-dessus de zéro. Cela indique que les estimations de C2 sont supérieures à celles de C1. La dispersion limitée et sa proximité de zéro montrent que, même si les estimations de C2 sont supérieures, les écarts entre C1 et C2 sont minimes et constants dans toutes les provinces.

Pour le groupe du SCIAN 7224, bien que tous les points soient dispersés également autour de zéro, leur dispersion est grande. Cela suppose que les écarts entre les estimations de C1 et C2 dans ce groupe sont grands pour quelques-unes des provinces.



Section 3 : Méthodologie utilisée pour établir des liens entre les anciennes et les nouvelles séries

Les séries C1 et C2 ont été combinées, à l'aide d'un processus d'étalonnage. Comme il est indiqué ci-dessus, les mouvements mensuels de C1 et C2 sont très similaires. Par conséquent, il a été décidé de baser les séries raccordées sur les changements mensuels des séries C1. (Remarque : Si le niveau des séries C1 et C2 avait aussi été très similaire,

les séries C2 auraient pu simplement continuer à partir des séries C1 sans avoir besoin d'une méthode de raccordement.)

Avant le raccordement, nous avons procédé à des « révisions annuelles » des séries C1. Les révisions annuelles sont effectuées une fois par an et sont fondées sur des données à jour et plus complètes des fichiers de TPS, des données fournies par les déclarants en retard et des révisions apportées aux données recueillies. Par conséquent, les séries C1 dans l'ensemble de données raccordées, sont les séries révisées annuellement pour C1.

L'objectif de la méthode de raccordement est de produire un seul ensemble de séries révisées pour la période de septembre 2007 à mars 2009 (mars a été choisi comme mois de raccordement après l'essai parallèle de six mois de C1 et C2). Pour les données antérieures à septembre 2007, nous utilisons les séries C1 (révisées annuellement). Après mars 2009, nous utilisons les séries C2. Pour la période située entre septembre 2007 et mars 2009, nous utilisons des méthodes statistiques pour établir des liens entre les deux séries.

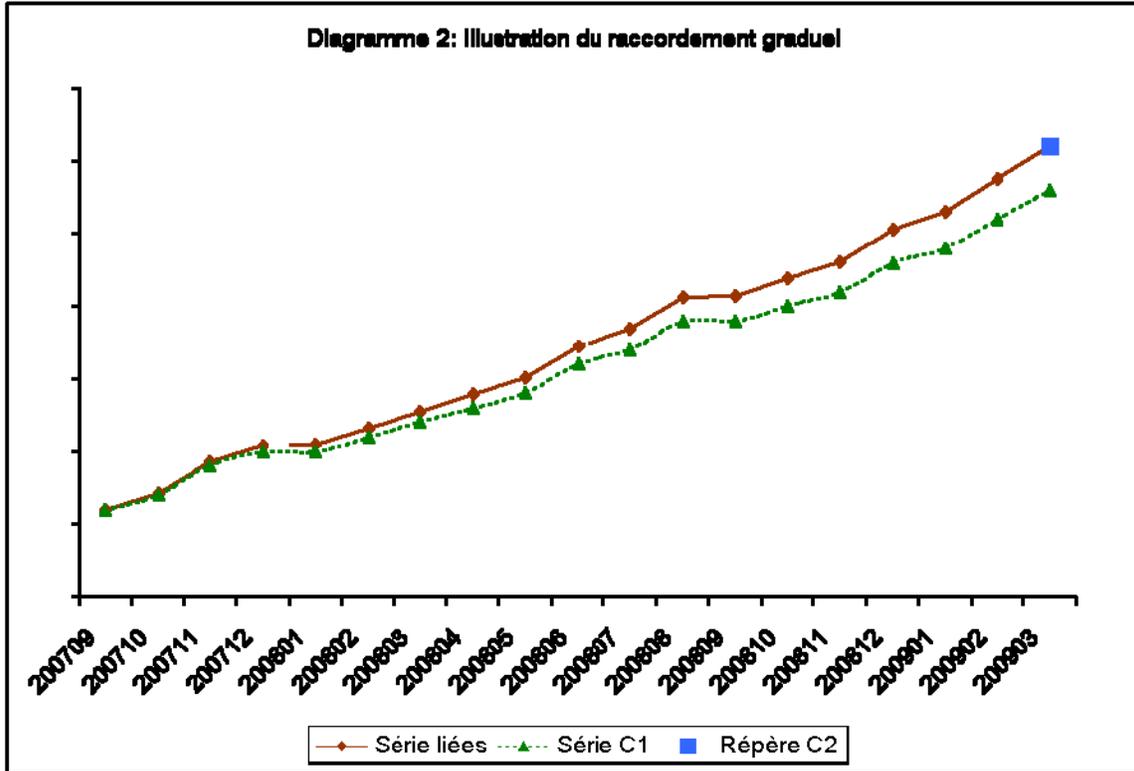
Il existe deux méthodes de raccordement : la méthode de raccordement graduel et la méthode de raccordement avec changement (brusque) de niveau. Quand les écarts entre les estimations de C1 et C2 étaient relativement petits, nous avons utilisé la méthode de raccordement graduel. Quand les écarts étaient plus grands, nous avons utilisé la méthode de raccordement avec changement de niveau.

Nous avons procédé au raccordement selon les provinces et selon les codes à 4 chiffres du SCIAN. Environ la moitié des domaines utilisaient la méthode de graduelle (ce qui représente environ 70 % du produit des ventes de l'EMSRDB), tandis que l'autre moitié utilisait la méthode de raccordement avec changement de niveau.

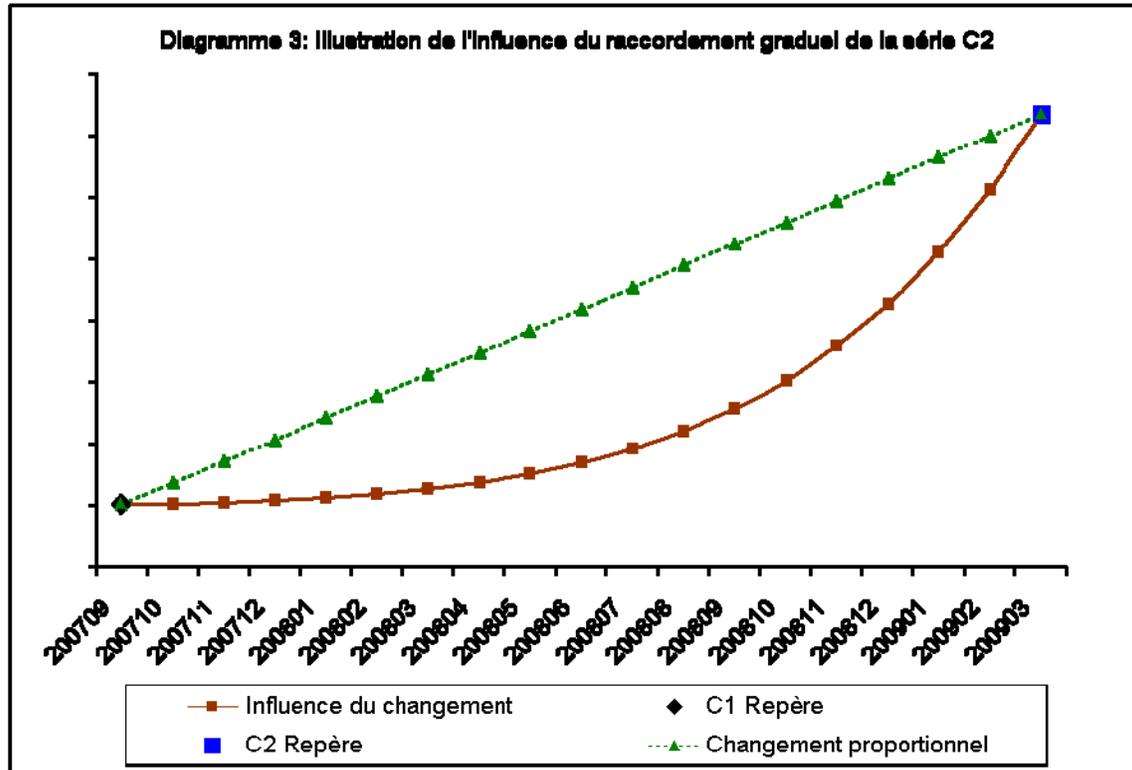
i) Méthode de raccordement graduel

Dans cette approche, la série débute en septembre 2007, avec l'estimation de C1. C'est le repère de départ (également appelé point d'ancrage). L'estimation de C2 au mois de mars 2009 est le repère de fin. La différence entre les estimations de C1 et C2 au mois de mars 2009 représente le montant raccordé graduellement dans les séries C1, rétrospectivement jusqu'à septembre 2007.

Le diagramme 2 illustre la méthode de raccordement graduel : une méthode qui préserve l'estimation C1 de départ et l'estimation C2 de fin et qui dérive vers une série raccordée entre les deux. La ligne pointillée représente la série initiale C1 et la ligne unie représente les nouvelles séries raccordées.



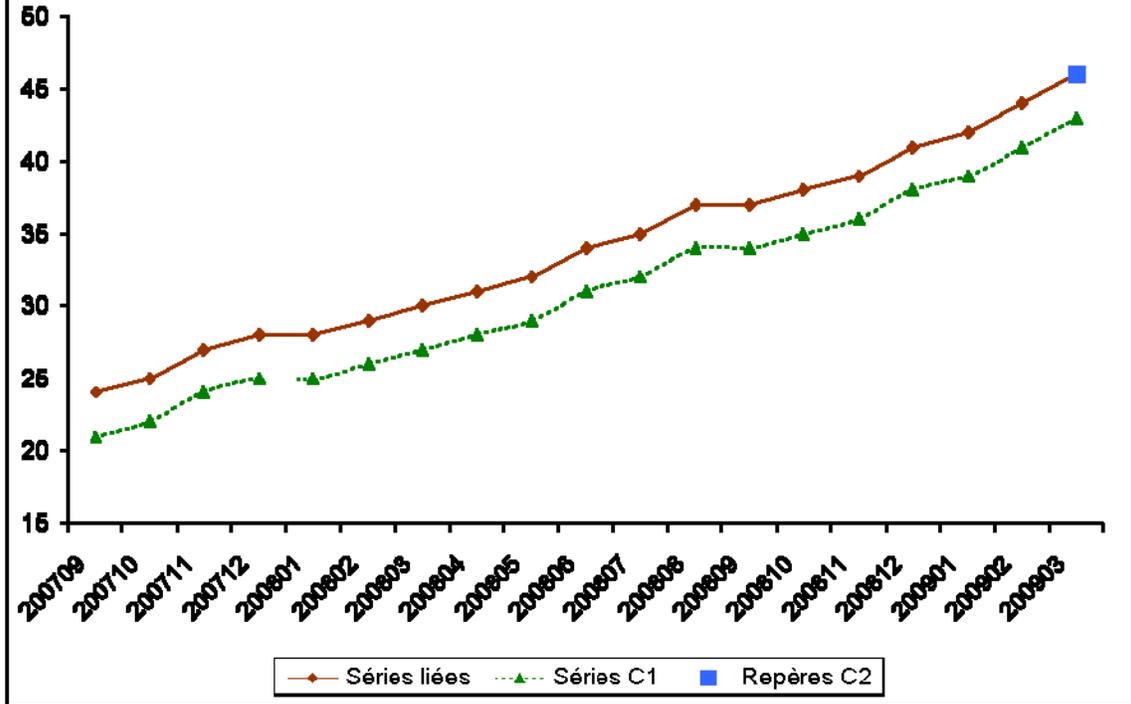
Dans le diagramme 3, la ligne unie illustre la façon dont est raccordée graduellement la différence entre les estimations de C1 et C2. Elle montre l'influence relative de C2 au fil du temps. Au début des séries, le repère C2 a peu d'influence et la série C1 domine. Au fur à mesure que les mois passent, le repère C2 de mars 2009 domine. Il tire la série C1 vers lui. La ligne pointillée, bien que non utilisée, démontre une répartition uniforme de la différence entre C1 et C2 pendant les 18 mois.



ii) Méthode de raccordement avec changement brusque de niveau

Dans cette approche, il y a seulement un repère, l'estimation de C2 du mois de mars 2009. À partir de cette estimation de C2, la série raccordée suit à rebours la tendance mensuelle de la série C1. La distance en pourcentage entre la série raccordée et la série C1 est constante et est égale à la variation de pourcentage en mars 2009, $\{100 * (C2 - C1) / C1\}$. Ainsi, dans le diagramme 4, la série raccordée, représentée par une ligne unie, est « parallèle » à la série C1, la ligne pointillée.

Diagramme 4: Illustration de raccordement avec changement brusque de niveau



Annexe 1 : Glossaire

Biais	Une mesure de la différence entre l'attente d'un estimateur et la valeur réelle du paramètre évalué.
EMSRDB	Enquête mensuelle sur les services de restauration et débits de boissons
Entreprises et établissements	<p>Les entreprises et les établissements sont les représentations statistiques d'une entreprise utilisée pour élaborer les plans d'échantillonnage.</p> <p>Dans le Registre des entreprises, si la structure opérationnelle d'une entreprise a des entités commerciales « enfants », l'entreprise est alors appelée entreprise « complexe ». L'entreprise au sommet de la structure opérationnelle est appelée entreprise et ses « enfants » sont appelés établissements.</p> <p>Au Canada, dans l'économie globale, 99,1 % des entreprises sont « simples ». Cela signifie qu'une seule entité représente la structure opérationnelle et la représentation statistique dans le RE.</p> <p>En effet, l'entreprise statistique et l'établissement statistique ne font qu'un!</p> <p>Les entreprises complexes, même si elles sont peu nombreuses (0,9 %), représentent 60 % de l'économie.</p>
Erreur type	Une mesure de la variabilité d'échantillonnage ou de la précision de l'estimation.
RE	Le Registre d'entreprises est une liste de toutes les entreprises au Canada provenant de Statistique Canada et tenue à jour par Statistique Canada. On le mentionne aussi comme la « base » dont provient un plan d'échantillonnage et à partir duquel on sélectionne un échantillon.
SCIAN	Système de Classification des Industries de l'Amérique du Nord (SCIAN). Les entreprises et les établissements appartenant à la population de l'EMSRDB commencent avec le code SCIAN 722.
TPS	Taxe sur les produits et services : les données de la TPS sont des données administratives auxquelles Statistique Canada accède par l'intermédiaire de l'Agence du revenu du Canada (ARC). Ces données fournissent les recettes mensuelles et sont utilisées pour remplacer les données d'enquête recueillies, ce qui réduit le fardeau

	de réponse et les coûts.
Types de strates de plan d'échantillonnage	<p><u>Tirage complet</u> (TC) : Toutes les unités dans cette catégorie sont sélectionnées pour la collecte de données en fonction de la taille ou d'un autre critère prédéfini. Ce sont les contributeurs les plus importants dans une industrie et ils ne peuvent pas être représentés par d'autres unités dans la population; en fait, chaque unité s'auto représente.</p> <p><u>Tirage partiel</u> (TP) : Un échantillon d'unités est sélectionné pour la collecte de données et chaque unité sélectionnée est pondérée pour représenter les autres unités qui n'ont pas été sélectionnées.</p> <p><u>Tirage nul</u> (TN) : Aucune unité n'est sélectionnée pour la collecte de données. Au lieu de cela, nous utilisons les données administratives pour ces unités.</p>
Unité d'échantillonnage	L'unité d'échantillonnage de l'EMSRDB est l'entreprise. Il s'agit d'une grappe d'établissements ayant le code SCIAN 722 et qui appartient à l'entreprise (enfant). L'établissement dominant est l'établissement avec la plus grande taille des recettes annuelles dans le RE. Son code SCIAN (cinq chiffres) et sa province sont le code SCIAN / Province assigné à la grappe (entreprise).

Annexe 2 : Aperçu de l'EMSRDB : but, objectifs, défis et méthodologie

But :

L'EMSRDB publie les estimations de la valeur mensuelle des ventes et le nombre d'emplacements par province ou territoire pour les industries de la restauration basées sur le Système de Classification des Industries de l'Amérique du Nord (SCIAN). Les classes à 4 chiffres du SCIAN qui font partie du champ de l'enquête sont indiquées ci-dessous :

SCIAN 7221	Restaurants à service complet
SCIAN 7222	Établissements de restauration à service restreint
SCIAN 7223	Traiteurs, entrepreneurs en restauration, cantines et comptoirs mobiles
SCIAN 7224	Débites de boissons

L'EMSRDB est une refonte de l'Enquête mensuelle sur les restaurants, traiteurs et tavernes (EMRTT), qu'elle a remplacée. On a procédé à la refonte en 2006 et 2007 et, le 1^{er} février 2008, les estimations préliminaires du mois de référence de novembre 2007 ont été publiées dans *Le Quotidien* de Statistique Canada. Le mois de référence de septembre 2007 a été utilisé pour faire le lien entre la série originale de l'EMRTT et les données remaniées de l'EMRDB en utilisant la méthode de raccordement avec changement brusque de niveau.

Objectifs :

Les objectifs de la refonte initiale consistaient à :

- améliorer la population cible afin de couvrir les entreprises sans salariés (en plus des employeurs de l'ancienne EMRTT);
- réduire le fardeau de réponse et faire une utilisation plus judicieuse des ressources de collecte en :
 - mettant en œuvre davantage d'ententes de collecte de données pour les franchiseurs;
 - introduisant des strates à tirage nul;
 - élargissant l'utilisation des données administratives (TPS);
- produisant des données désaisonnalisées;
- procédant périodiquement à une restratification pour mettre à jour le plan d'échantillonnage et améliorer les estimations de l'EMSRDB.

Défis :

Il y a plusieurs défis à relever dans l'EMSRDB en ce qui a trait à l'estimation. Premièrement, la population de ces secteurs est extrêmement volatile en ce qui concerne les naissances, les fermetures et les changements de propriétaire. Il est important de les identifier tous les mois, pour s'assurer que les changements dans la population apparaissent dans les estimations de l'échantillon. Deuxièmement, les commerces à succursales influencent considérablement les estimations. L'identification et la classification des emplacements des commerces à succursales, contribuent à améliorer la qualité de la collecte de données en mettant au point des modalités de déclaration qui réduisent considérablement le fardeau de réponse, tout en augmentant la couverture des unités dans la population.

Au cours d'un mois, quatre ensembles d'estimations sont publiés : l'estimation préliminaire du dernier mois de référence, 'm', la première révision du mois de référence précédente, 'm-1', et la deuxième révision du mois de référence, 'm-2'. De plus, les estimations désaisonnalisées font l'objet d'un mois supplémentaire de révision pour la troisième révision du mois de référence, 'm-3'. Il y a une exception durant la publication du mois de référence de mars, lorsqu'on révisé rétrospectivement les données désaisonnalisées jusqu'au mois de janvier de l'année précédente.

Méthodologie :

a) Base et unité de sondage :

La population admissible à l'EMSRDB est une liste d'établissements ayant le code SCIAN 722, provenant du Registre des entreprises (RE). L'unité d'échantillonnage de l'EMSRDB est l'entreprise (voir le glossaire de l'annexe 1).

Dans l'EMSRDB, pour un mois donné, environ 1,7 % des entreprises (unités d'échantillonnage) sont « complexes » (c'est-à-dire que l'entreprise compte plus d'un établissement dans différentes provinces ou codes SCIAN) et 98,3 % des entreprises sont « simples » (c'est-à-dire que l'entreprise compte un établissement ou plus dans la même province et le même code SCIAN).

b) Plan d'échantillonnage, estimation et utilisation des données sur la TPS :

Un objectif de la refonte de l'EMSRDB était d'augmenter l'utilisation des données administratives. Les données de la taxe sur les produits et services (TPS) de l'Agence du revenu du Canada fournissent les recettes mensuelles. L'utilisation des données de la TPS (à la place des données d'enquête recueillies) vise à réduire les coûts de collecte des données ainsi que le fardeau de réponse et à améliorer la qualité des données.

On peut utiliser les données de la TPS pour les entreprises simples, car il y a un lien individuel entre l'unité de l'établissement dans l'EMSRDB et ses recettes mensuelles basées sur la TPS. Cependant, le lien pour les entreprises complexes avec les données de la TPS est univoque ou multivoque. Par conséquent, nous n'utilisons pas les données de la TPS pour ces cas. Nous devons donc obtenir les données sur les ventes mensuelles par le biais des données d'enquêtes recueillies pour ces unités complexes.

Une limite des données de la TPS est qu'il nous faut souvent attendre plusieurs mois avant de connaître la date à laquelle des unités ont cessé leurs activités. Afin d'en tenir compte, on applique un taux de fermeture pour ajuster le nombre prévu de fermetures pendant le mois de référence. Au moment de la révision annuelle, nous disposons d'informations empiriques pour plusieurs mois sur les fermetures réelles, que nous pouvons appliquer aux estimations historiques.

c) Plan d'échantillonnage :

Il y a deux méthodes d'estimation dans l'EMSRDB. Le plan d'échantillonnage traditionnel basé sur les données d'enquête recueillies dans le cadre de l'EMSRDB en est une. Le modèle pour les entreprises simples, basé sur les données de la TPS, est la deuxième méthode utilisée pour certaines provinces et certains domaines du SCIAN.

i) Plan traditionnel :

Dans chaque province et chaque domaine du SCIAN, il y a ; une strate à tirage complet (dans laquelle toutes les unités d'une strate sont sélectionnées), deux strates à tirage partiel (un échantillon aléatoire d'unités est sélectionné) et une strate à tirage nul (aucune unité n'est sélectionnée et on utilise à la place les données de la TPS). Les grandes unités d'échantillonnage sont attribuées à la strate à tirage complet; les unités d'échantillonnage de taille moyenne ou petite sont attribuées à la strate à tirage partiel et les entreprises simples les plus petites sont dans la strate à tirage nul.

Les limites entre ces strates (en particulier la cellule SCIAN/province) sont établies en se basant sur la mesure des recettes annuelles de l'EMSRDB à partir de la taille des entreprises figurant dans le RE à l'aide d'un algorithme (Hidiroglou-Berthelot) conçu pour dériver la taille minimale de l'échantillon requis pour un niveau d'exactitude souhaité.

Pour la strate à tirage partiel, on utilise l'échantillonnage aléatoire simple sans remise (EASSR). Ainsi, dans une strate toutes les unités d'échantillonnage ont la même probabilité d'être sélectionnées. Pour une strate, partons de l'hypothèse que « N » est le nombre d'unités de la base de sondage et « n » le nombre d'unités sélectionnées dans l'échantillon. Par conséquent, la probabilité de chaque unité d'être sélectionnée est « n/N ». Pour l'estimation, on utilise le même estimateur Horvitz-Thompson : la pondération attribuée à une unité échantillonnée est donc $\{1/(n/N)\} = N/n$. La pondération attribuée à chaque unité de la strate à tirage complet est 1 (la pondération des unités à succursales est également 1).

- On donne à chaque naissance d'entreprise une probabilité d'être sélectionnée ($P_{\text{naissance}}$) dans l'échantillon. $P_{\text{naissance}}$ est égal à la probabilité utilisée pour sélectionner l'échantillon initial dans la strate. Il est possible qu'aucune, que certaines ou que toutes les créations d'un mois soient sélectionnées, mais nous prévoyons que le nombre de naissances sélectionnées sera près de (B multiplié par $P_{\text{naissance}}$). La pondération de l'échantillonnage de la strate passe de N/n à $\{(N+B)/(n+b)\}$ où B est le nombre de naissances détectées dans le RE et « b » le nombre de naissances sélectionnées dans l'échantillon.

La source des données pour le plan d'échantillonnage traditionnel est les données d'enquête de l'EMSRDB pour les entreprises avec succursales, les unités des strates à tirage complet et des strates à tirage partiel. Les données sur les recettes des fichiers de la TPS sont la source des données pour la strate à tirage nul, obtenues à l'aide du modèle pour les entreprises simples décrites ci-dessous.

ii) Modèle pour les entreprises simples :

Le modèle pour les entreprises simples de l'EMSRDB est fondé sur l'estimation des ventes totales de l'EMSRDB à l'aide de l'estimateur par ratio.

Nous sélectionnons un petit échantillon d'entreprises simples et nous recueillons leurs réponses sur les ventes à l'EMSRDB. Nous relierons cette unité échantillonnée à l'unité de TPS qui lui est lié pour obtenir les recettes de la TPS. Puis le ratio ; (la somme des données sur les ventes recueillies par l'EMSRDB)/(la somme des recettes de la TPS) est calculé pour chaque province par cellule du groupe SCIAN. Ce ratio est celui utilisé dans le modèle pour les entreprises simples de l'EMSRDB. Nous l'utilisons pour chaque strate à tirage nul et pour une partie de la population de l'EMSRDB (voir le diagramme 5).

Le but du ratio est double. Premièrement, estimer (ou modéliser) le lien conceptuel entre les données sur les ventes recueillies par l'EMSRDB et les données sur les recettes de la

TPS. Habituellement, ce ratio est près de 1, ce qui laisse supposer qu'il y a peu de différence conceptuelle entre une réponse sur les ventes à l'EMSRDB et une réponse sur les recettes à la TPS.

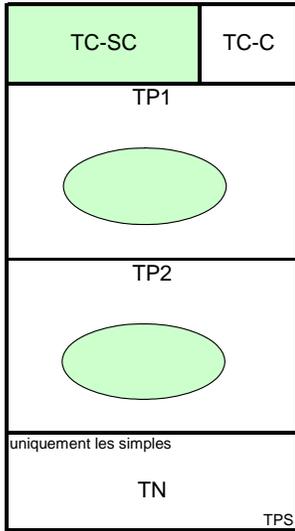
Deuxièmement, les données de la TPS du mois de référence ne sont pas encore disponibles pour l'estimation préliminaire. Au lieu de cela, nous utilisons les données de la TPS du mois précédent et, par conséquent, le modèle procède aussi à une estimation du changement économique de « m-1 » à « m ».

Le diagramme 5, ci-dessous, résume la base d'échantillonnage de l'EMSRDB.

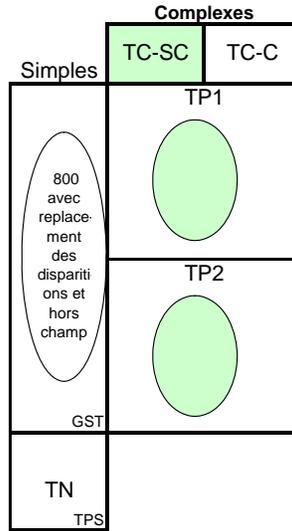
1. On utilise le plan d'échantillonnage traditionnel avec les strates à tirage complet, à tirage partiel 1 et 2 et à tirage nul pour :
 - Terre-Neuve-et-Labrador (10), l'Île-du-Prince-Édouard (11) la Nouvelle-Écosse (12), et le Nouveau-Brunswick (13) dans les groupes SCIAN 7221, 7222, 7223 et 7224
 - le Québec (24), l'Ontario (35), le Manitoba (46), la Saskatchewan (47), l'Alberta (48) et la Colombie-Britannique (59) dans les groupes SCIAN 7223 et 7224.
2. Nous utilisons le plan d'échantillonnage traditionnel avec les strates à tirage complet, à tirage partiel 1 et 2 et à tirage nul (TN) pour les entreprises complexes dans les provinces 24, 35, 46, 47, 48 et 59 et dans les groupes SCIAN 7221 et 7222.
3. Pour les entreprises simples, on utilise le modèle pour les entreprises simples dans les provinces 24, 35, 46, 47, 48 et 59 et dans les groupes SCIAN 7221 et 7222.
4. Le Yukon (60), les Territoires-du-Nord-Ouest (61) et le Nunavut (62) sont limités à deux strates : une strate à tirage complet et une strate à tirage nul.

Diagramme 5 : La conception de l'EMSRDB

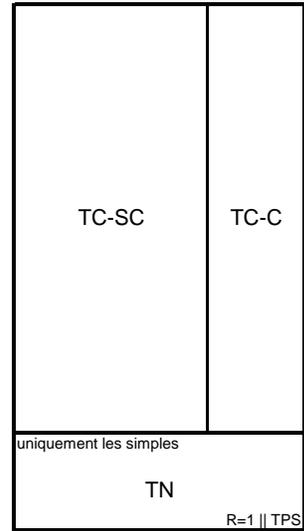
Province: 10-13 || SCIAN: 7221-7224
 Province: 24-59 || SCIAN: 7223-7224



Province: 24-59 || SCIAN: 7221-7222



Province: 60-62 || SCIAN: 7221-7224



Échantillonnage classique

TC-SC = Strate tirage complet sans les chaînes || TC-C = Strate tirage complet avec chaînes
 TP1 = Strate tirage partiel 1 || TP2 = Strate tirage partiel 2 || TN = Strate à tirage nul